



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/720,798	11/24/2003	Kirill Stoimenov	9432-000248	5424

27572 7590 07/25/2006

HARNESSE, DICKEY & PIERCE, P.L.C.  
P.O. BOX 828  
BLOOMFIELD HILLS, MI 48303

EXAMINER

WOZNIAK, JAMES S

ART UNIT PAPER NUMBER

2626

DATE MAILED: 07/25/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

**Office Action Summary**

Application No.

10/720,798

Applicant(s)

STOIMENOV ET AL.

Examiner

James S. Wozniak

Art Unit

2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 24 November 2003.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-46 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-46 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 24 November 2003 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- ☒ Notice of References Cited (PTO-892)
- ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)  
Paper No(s)/Mail Date \_\_\_\_\_
- ☐ Interview Summary (PTO-413)  
Paper No(s)/Mail Date. \_\_\_\_\_
- ☐ Notice of Informal Patent Application (PTO-152)
- ☐ Other: \_\_\_\_\_

## DETAILED ACTION

### *Claim Objections*

1. **Claims 10 and 34** are objected to because of the following informalities:

In claims 10 and 34, Line 2, "on" should be changed to --in--.

Appropriate correction is required.

### *Claim Rejections - 35 USC § 103*

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

3. **Claims 1-4, 10-11, 27-28, 30, and 34-35** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al (*U.S. Patent: 6,990,448*) in view of Baker et al (*U.S. Patent: 6,092,044*).

With respect to **Claims 1 and 27**, Charlesworth discloses a voice annotation (tag) system, that allows a user to generate text annotations corresponding to a media file through speech-to-text conversion, wherein the annotations (tags) comprise text (*alphanumeric characters indicative of a voice tag*) and an associated phoneme string (normalized text that serves as recognition text during retrieval) (*Col. 9, Line 61- Col. 10, Line 30*).

Charlesworth does not teach how a voice recognizer utilized in voice tag generation is trained to identify spoken words corresponding to voice tags, specifically using the means noted in claims 1 and 27, however Baker discloses:

An editor receptive of alphanumeric characters indicative of a word, the editor configured to display and edit the alphanumeric characters (*word editor accepting typed and speech-generated text, Fig. 17, Col. 17, Line 66- Col. 18, Line 6*);

A text parser connected to the editor and operable to generate normalized text corresponding to the alphanumeric characters, such that the normalized text serves as recognition text for the word and is displayed by the editor (*segmenting a recognized word into a phoneme sequence, Col. 15, Line 56- Col. 16, Line 5; and editor display, Fig. 17*); and

A storage mechanism connected to the editor and operable to update a lexicon with the displayed alphanumeric characters and the corresponding normalized text, thereby developing a "sounds like" pair (*dictionary storage of word and phonetic spelling (pronunciation) pairs, Col. 16, Lines 1-5*).

Charlesworth and Baker are analogous art because they are from a similar field of endeavor in word recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth with the speech recognition word editor taught by Baker in order to allow a user to add and personalize words (corresponding to voice tags in the case of Charlesworth) in a speech recognition vocabulary by representing how a word is spoken by a particular speaker (*Baker, Col. 1, Lines 9-21 and 50-57*).

With respect to **Claim 2**, Baker further recites:

The alphanumeric characters are typed in via a keyboard connected to the editor  
*(keyboard, Fig. 1, Element 115).*

Also, Charlesworth discloses the use of a keyboard for entering a voice annotation *(Col. 10, Lines 36- 53).*

With respect to **Claims 3 and 28**, Charlesworth recites the voice annotations as applied to claim 1, while Baker further discloses:

The word editor is connected to the lexicon *(editor connected to a dictionary for word addition, Col. 15, Line 56- Col. 16, Line 6)* and further configured to display a list of words residing in the lexicon *(displaying word history, Col. 18, Lines 52-58).*

With respect to **Claim 4**, Charlesworth discloses a phoneme symbol sequence (normalized text) used by a speech recognizer in a voice annotation (tagging) system (Col. 9, Line 61- Col. 10, Line 30), while Baker recites the dictionary containing phoneme sequence data for use by a speech recognizer *(Fig. 2, Elements 230 and 245)* as applied to Claim 1.

With respect to **Claims 10 and 34**, Baker further discloses:

The editor is configured to modify existing word "sounds like" pairs stored in the lexicon *(editing a word history using an editor, Col. 18, Lines 42-58).*

With respect to **Claims 11 and 35**, Baker further recites:

The editor is configured to modify a phonetic transcription used by the speech recognizer *(editing a word pronunciation comprising a phoneme sequence transcription, Col. 17, Lines 66- Col. 18, Line 6; and Col. 18, Lines 42-58).*

With respect to **Claim 30**, Baker teaches the phoneme sequence utilized in speech recognition as applied to Claim 27.

4. **Claims 5 and 31** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al in view of Baker et al, and further in view of Hashimoto et al (U.S. Patent: 5,632,002).

With respect to **Claims 5 and 31**, Charlesworth in view of Baker discloses the voice annotation dictionary editor display as applied to Claims 1 and 27. Although Charlesworth discloses the use of topic-based dictionaries (*Col. 6, Lines 15-22*), Charlesworth in view of Baker does not specifically suggest displaying a dictionary topic, however Hashimoto recites a means for displaying speech recognition dictionary names indicative of dictionary content (*Col. 68, Lines 50-65*).

Charlesworth, Baker, and Hashimoto are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth in view of Baker with the dictionary content display means taught by Hashimoto in order to allow a user to easily access various dictionary contents for editing operations that can reduce memory requirements (*Hashimoto, Col. 68, Lines 31-42*).

5. **Claims 6-7, 9, 17-18, 32, and 41** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al in view of Baker et al, and further in view of Morrison (U.S. Patent: 5,425,128).

With respect to **Claims 6 and 32**, Charlesworth in view of Baker discloses the voice annotation dictionary editor as applied to Claims 1 and 27. Although Charlesworth discloses that

dictionary words may be imported from an input file (Col. 18, Lines 42-58), Charlesworth in view of Baker does not specifically suggest importing a lexicon from an external data source, however Morrison discloses importing a vocabulary for a particular speech recognition application from an external host computer (*Col. 5, Lines 25-60*).

Charlesworth, Baker, and Morrison are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth in view of Baker with the vocabulary importing means disclosed by Morrison in order to implement the ability to process speech from a particular speaker at multiple computer systems without the need for portable media (*Morrison, Col. 2, Lines 30-51*).

With respect to **Claim 7**, Charlesworth recites the voice annotation system as applied to claim 1, while Morrison further recites:

The external data source receives a request from a client computer (*vocabulary request based on a particular application, Col. 5, Lines 25-60*).

With respect to **Claim 9**, Morrison further recites:

The request includes content requirements for a lexicon (*vocabulary request based on a particular application, Col. 5, Lines 25-60*).

With respect to **Claims 17-18 and 41**, Morrison further recites:

Uploading the lexicon including a content description to a remote location (*uploading a speech recognition vocabulary, associated with a particular application (content description), to a host computer Col. 9, Lines 46-48*).

6. **Claims 8 and 33** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al in view of Baker et al, further in view of Morrison, and yet further in view of Hashimoto et al.

With respect to **Claims 8 and 33**, Charlesworth et al in view of Baker et al, and further in view of Morrison discloses the voice annotation dictionary that can be downloaded from a host computer as applied to Claims 6 and 32. Charlesworth et al in view of Baker et al, and further in view of Morrison do not specifically suggest displaying an available dictionary list, however Hashimoto discloses displaying such a list (*Col. 21, Lines 25-40*).

Charlesworth, Baker, and Morrison are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth in view of Baker with the vocabulary importing means disclosed by Morrison in order to enable a user to quickly switch to an appropriate vocabulary and suppress the recognition error rate (*Hashimoto, Col. 21, Lines 25-40*).

7. **Claims 12 and 36** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al in view of Baker et al, and further in view of Korall et al (*U.S. Patent: 6,996,531*).

With respect to **Claims 12 and 36**, Charlesworth in view of Baker discloses the voice annotation dictionary editor utilizing text-to-phoneme conversion, as applied to Claims 1 and 27. Charlesworth in view of Baker does not specifically suggest that a user is prompted is



normalized text cannot be generated by a parser, however Korall recites prompting a speaker if an input cannot be resolved into phonemes (*Col. 8, Lines 31-38*).

Charlesworth, Baker, and Korall are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth in view of Baker with the prompting means disclosed by Korall in order to provide an effective recognition fallback method if an input cannot be resolved (*Korall, Col. 8, Lines 31-38*).

8. **Claims 13-16, 37-40** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al in view of Baker et al, and further in view of Young et al (*U.S. Patent: 6,064,959*).

With respect to **Claims 13 and 37**, Charlesworth in view of Baker discloses the voice annotation dictionary editor utilizing text-to-phoneme conversion, as applied to Claims 1 and 27. Charlesworth in view of Baker do not specifically suggest speech recognition data verification, however Young recites a means for verifying and correcting a speech recognition word pronunciation (*Col. 21, Lines 35-61*).

Charlesworth, Baker, and Young are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth in view of Baker with the means for verifying and correcting a speech recognition word pronunciation as taught by Young in order to provide a computer-implemented means for speech recognition error correction (*Young, Abstract*).

With respect to **Claims 14 and 38**, Young further discloses correcting (modifying) a pronunciation corresponding to a particular word (*Col. 21, Lines 11-26*).

With respect to **Claims 15 and 40**, Young further discloses the use of an n-best list (*Col. 21, Lines 35-61*).

With respect to **Claim 16**, Young further discloses the use of a recognition hypothesis score (*Col. 4, Lines 34-51*).

**Claim 39** contains subject matter similar to Claims 11 and 14, and thus, is rejected for the same reasons.

9. **Claims 19-22, 29, 42, and 43** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al in view of Baker et al, and further in view of Sabourin et al (*U.S. Patent: 6,073,099*).

With respect to **Claims 19 and 42**, Charlesworth in view of Baker discloses the voice annotation dictionary editor utilizing text-to-phoneme conversion, as applied to Claims 1 and 29. Charlesworth in view of Baker do not specifically suggest identifying recognition text that is confusingly similar, however Sabourin recites a means for identifying a confusability cost between two phonetic transcriptions (recognition text) (*Col. 3, Line 31- Col. 4, Lines 13*).

Charlesworth, Baker, and Sabourin are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth in view of Baker with the means for identifying a confusability cost between two phonetic transcriptions as taught by Sabourin in order to provide a means for automatically generating an objective

Art Unit: 2626

metric of the likelihood of confusing a spoken word with another spoken word (*Sabourin, Col. 1, Lines 31-34*).

With respect to **Claims 20 and 43**, Sabourin further discloses:

Identifying the at least one other word includes calculating a measure distance between phonetic transcriptions associated with each recognition text, where the measure distance is indicative of similarity between the phonetic transcriptions (*calculation of a Levenshtein distance between two phonetic transcriptions, Col. 4, Lines 15-50; and Col. 1, Lines 40-51*).

With respect to **Claim 21**, Sabourin further recites:

The measure distance is based on a number of edit operations needed to make the phonetic transcriptions identical (*Levenshtein distance based upon editing operations, Col. 4, Lines 15-50*).

With respect to **Claim 22**, Sabourin further recites:

Providing alternative recognition text of the desired word (*replacing confusable words with alternative non-confusable synonyms, Col. 10, Line 60- Col. 11, Line 8*).

**Claim 29** contains subject matter similar to Claims 19 and 22, and thus, is rejected for the same reasons.

10. **Claims 23-24, 29, and 44** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al in view of Baker et al, and further in view of Hirayama et al (*U.S. Patent: 6,708,150*).

With respect to **Claims 23, 29, and 44**, Charlesworth in view of Baker discloses the voice annotation dictionary editor utilizing text-to-phoneme conversion, as applied to Claim 1.

Charlesworth in view of Baker do not specifically suggest identifying an unbalanced phrase length, however Hirayama discloses a means for detecting phrases exceeding a specific length (*Col. 12, Line 63- Col. 13, Line 15*).

Charlesworth, Baker, and Hirayama are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth in view of Baker with the means for detecting phrases exceeding a specific length as taught by Hirayama in order to increase speech recognition reliability by utilizing shorter word alternatives (*Hirayama, Col. 12, Line 47- Col. 13, Line 15*).

With respect to **Claim 24**, Hirayama discloses the use of shorter word alternatives as applied to Claim 23.

11. **Claims 25-26, 29, and 45-46** are rejected under 35 U.S.C. 103(a) as being unpatentable over Charlesworth et al in view of Baker et al, and further in view of Goronzy (*U.S. Patent Publication: 2002/0111805*).

With respect to **Claims 25, 29, and 45**, Charlesworth in view of Baker discloses the voice annotation dictionary editor utilizing text-to-phoneme conversion, as applied to Claims 1 and 27. Charlesworth in view of Baker do not specifically suggest identifying words that are hard to pronounce, however Goronzy recites detecting words that have difficult pronunciations (*Paragraphs 0005 and 0053*).

Charlesworth, Baker, and Goronzy are analogous art because they are from a similar field of endeavor in speech recognition. Thus, it would have been obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings of Charlesworth in view of Baker with the means for detecting words that have difficult pronunciations as taught by Goronzy in order to manage the problem of decreasing recognition rates for speech in a target language given by a non-native speaker (*Goronzy, Paragraph 0005*).

With respect to **Claims 26 and 46**, Goronzy further recites:

Providing alternative recognition text of the desired word (*providing pronunciation variants for particular words, Paragraph 0058*).

### *Conclusion*

12. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure:

Morag (*U.S. Patent: 6,324,585*)- teaches a means for adding a voice annotation to images or video.

Berstis (*U.S. Patent: 6,721,001*)- teaches an annotation system utilizing speech-to-text conversion.


Tahara et al (*U.S. Patents: 6,952,675 and 6,983,248*)- teaches a speech recognition method utilizing “sounds-like” spelling.

13. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James S. Wozniak whose telephone number is (571) 272-7632. The examiner can normally be reached on M-Th, 7:30-5:00, F, 7:30-4, Off Alternate Fridays.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached at (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

James S. Wozniak  
7/17/2006



**DAVID HUDSPETH**  
**SUPERVISORY PATENT EXAMINER**  
**TECHNOLOGY CENTER 2626**